



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/688,041	10/17/2003	Philip Gleason	BOC9-2003-0021 (390)	9201
40987	7590	09/21/2007	EXAMINER	
AKERMAN SENTERFITT			NG, EUNICE	
P. O. BOX 3188			ART UNIT	PAPER NUMBER
WEST PALM BEACH, FL 33402-3188			2626	
			MAIL DATE	DELIVERY MODE
			09/21/2007	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

<b>Office Action Summary</b>	<b>Application No.</b> 10/688,041	<b>Applicant(s)</b> GLEASON ET AL.	
	<b>Examiner</b> Eunice Ng	<b>Art Unit</b> 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 09 July 2007.
- 2a) ☒ This action is **FINAL**.                      2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-3,8-10,13-15,20-22,25 and 26 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-3,8-10,13-15,20-22,25 and 26 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- |  |   |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892)   | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)                       | 5) <input type="checkbox"/> Notice of Informal Patent Application                       |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)<br>Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____  |

## **DETAILED ACTION**

### ***Response to Amendment***

1. In response to the Office Action mailed 4/9/07, Applicants have submitted an Amendment, filed 7/9/07, canceling claims 4-7, 11, 12, 16-19, 23 and 24, amending claims 1, 8, 9, 13, 20, 21 and 25, adding new claim 26, without adding new matter, and arguing to traverse claim rejections.

### ***Response to Arguments***

2. Applicant's arguments with respect to claims 1-3, 8-10, 13-15, 20-22, 25 and 26 have been considered but are moot in view of the new ground(s) of rejection, below.

### ***Claim Rejections - 35 USC § 112***

3. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

4. Claims 9, 21 and 25, recite limitation(s) for which there is insufficient antecedent basis for the limitation(s) in the claim(s), as follows:

Claims 9 and 21 still recite the limitation "said automatically displaying parameters step" in lines 1-2 of the claims, and are dependent on claims 1 and 13, respectively, which only recite a step of "displaying parameters." The examiner has interpreted "said automatically displaying parameters step" to be --said displaying parameters step--. Appropriate correction is required.

Claim 25 recites "highlighting in the display of the original recording" in line 16 of the claim. However, there is no step for displaying the original recording in any of the previous

Art Unit: 2626

limitations. The examiner has interpreted there to be means for displaying the original recording containing a selected phonetic unit, similar to what is recited in independent claims 1 and 13, prior to the means for highlighting in claim 25. Appropriate correction is required.

### ***Claim Objections***

5. Claims 1, 2, 8, 9, 13, 14, 20 and 25 are objected to because of the following informalities:

In claim 1, line 12, and claim 13, line 13, “displaying original” should be --displaying an original--; and “containing selected phonetic” should be --containing a selected phonetic-- or --containing a user-selected phonetic--.

In claims 8 and 20, 4<sup>th</sup> line, “add a phonetic” should be --adding a phonetic--.

Claims 9 and 25 are missing a period at the end of the sentences. Claim 25, 15<sup>th</sup> line, “marker.” should be --marker;--.

Claims 2 and 14 recite “automatically displaying” in the 2<sup>nd</sup> line of the claims, but it is unclear what is meant by “automatically.” The examiner has interpreted “automatically displaying” to mean --displaying--.

Appropriate correction is required.

### ***Claim Rejections - 35 USC § 103***

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. Claims 1-3, 8-10, 13-15, 20-22, 25 and 26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Yamazaki, US Patent 5,864,814 in view of Miyatake *et al.* (hereinafter “Miyatake”), US Patent 5,842,167 and Yamazaki, US Patent 5,875,427 (hereinafter “Yamazaki2”), and further in view of Hejna, Jr., US Patent 7,043,433.

Regarding claims 1, 13 and 25, Yamazaki teaches a computer-implemented method for debugging and tuning synthesized audio (col. 22, line 28 – col. 26, line 28; Figs. 26-33), comprising the steps of: (c) displaying a waveform corresponding to the synthesized audio generated from concatenated phonetic units; and (e) displaying [an] original recording containing [a] selected phonetic unit; (col. 23, ll. 17-24, teaches “original waveform display window...synthesized waveform display window”); and (d) displaying parameters corresponding to at least one of the phonetic units (col. 23, ll. 41-49, “correlation between the parameters on the time axis”; col. 23, line 66 – col. 24, line 43, “pitch pattern W1...pitch pattern W2 of the synthesized waveform...pitch label”; col. 25, ll. 17-32, “velocity indicating a volume; col. 23, ll. 41-49, “correlation between the parameters on the time axis”; col. 25, ll. 61-67, “change of velocity...change of pitch...manual operation”; col. 26, ll. 23-28, “parameter”; Fig. 33 illustrates the pitch/velocity for a certain phoneme, which provides a visual indication of the pitch/velocity values (displaying parameters)).

Yamazaki teaches receiving voice-generating information (Abstract, col. 10, ll. 15-23). Yamazaki does not explicitly teach, but Miyatake teaches, (a) receiving a user-supplied text with a visual user interface; and (b) generating synthesized audio generated from concatenated phonetic units, the synthesized audio being a voice rendering of the user-supplied text; (col. 3, ll. 6-26, teaches “FIG. 1...inputting means...for inputting text data, a command and hand-written

Art Unit: 2626

characters...morpheme analyzing means 2...divide the text data into minimum language units each having the meaning...speech language processing means 4 determines synthesis units which are suitable for producing a sound from text data thereby to generate prosodic data”;

Abstract, “synthesizing speech from text data...text data displayed on a screen” and Figs. 1-4).

It would have been obvious for one of ordinary skill in the art at the time the invention was made to receive textual input and generate from the input synthesized audio as in Miyatake because text to speech synthesizing technology is readily available, and as Miyatake teaches in col. 2, ll. 32-34, this method also allows for operation in response to receiving hand-written text data.

Yamazaki teaches or suggests (d) the parameters including configuration parameters comprising a phonetic alignment marker, a phonetic unit label, and a pitch mark (col. 23, ll. 50-62, teaches, “phoneme is set in each of spaces separated from each other with a label...‘yo’...‘de’” (phonetic unit label); col. 24, ll. 21-43, and Fig. 33, elements D1, D3, D4, D5 teaches “pitch label,” the pitch label provides an indication of pitch on the display (pitch mark); phoneme labels are aligned with the waveform sections (phonetic alignment); col. 23, ll. 29-40, teaches in “step S104, to set a duration length of each phoneme in relation to the original waveform displayed on the original waveform display window 25B, labels [phonetic alignment markers] each separating phonemes from each other along the direction of a time axis are given through a manual operation”).

Yamazaki and Miyatake do not explicitly teach the parameters including configuration parameters comprising (d) a phonetic unit label electronically generated without additional user input. However, Yamazaki<sup>2</sup> teaches in col. 16, ll. 31-33, “it is possible to automate the steps from input of phoneme string information up to generation of a label by using the voice

Art Unit: 2626

recognition technology.” It would have been obvious for one of ordinary skill in the art at the time the invention was made to combine the teaching elements of Yamazaki and Miyatake with Yamazaki<sup>2</sup> because voice recognition technology is readily available and may be quicker than manual operation.

Yamazaki teaches: (f) receiving an editing input from the user (col. 25, line 61 – col. 26, line 16, “change of velocity...change of pitch...manual operation”); and (g) adjusting at least one configuration parameter in accordance with the editing input wherein adjusting includes repositioning the phonetic alignment marker (col. 23, ll. 29-40, teaches in “step S104, to set a duration length of each phoneme in relation to the original waveform displayed on the original waveform display window 25B, labels [phonetic alignment markers] each separating phonemes from each other along the direction of a time axis are given through a manual operation”; col. 26, ll. 3-5, teaches, “If change of a label is requested (step S114), system control returns to step S104, and the label [phonetic alignment marker] is changed [repositioned] through a manual operation”; see Fig. 24, elements S104 and S114).

Yamazaki teaches wherein the at least one configuration parameter is stored in a configuration file (Fig. 33; col. 25, ll.40-52, “new filing”; col. 25, line 61 – col. 26, line 28, “change of velocity...change of pitch...for each label”; col. 26, ll. 29-53, “change of parameters in...original synthesized waveform...object for editing”; col. 23, ll. 1-16, “making new voice-generating information”).

Yamazaki fails to teach, but Miyatake teaches, wherein the at least one configuration parameter is stored in a text-to-speech engine configuration file (col. 1, ll. 55-62, “receives text data and edition data...synthesizes speech corresponding to the character data”; col. 4, ll. 20-44,

Art Unit: 2626

“pauses are created”; Figs. 3-4; Miyatake teaches a visual speech synthesis editing system, like Yamazaki, and where it is used in a text-to-speech environment). Therefore, it would have been obvious to modify the teaching elements of Yamazaki with Miyatake in order to use editing data to produce a natural sounding spoken output from text, as described by Miyatake (col. 1, ll. 55-62 and ll. 20-30).

Yamazaki, Miyatake and Yamazaki<sup>2</sup> do not explicitly teach, but Hejna, Jr. suggests: (h) highlighting in the display of the original recording at least one user-selected phonetic unit (Fig. 25, ll. 15-66, teaches, “produce a two dimensional graph...time and possibly text or phonetic words, displayed on a horizontal axis...displays a two dimensional representation of the input MW (the input audio or audio-visual work) with text or phonetic labels...well known...phonetic information...displayed as an overlay on top of a graphical representation of a speech waveform...Audience member (user) can highlight regions of the text displayed on Graphical Display...to identify specific portions of the input MW that are associated with the highlighted text”; see also Figs. 5 and 6, particularly element 6100). It would have been obvious for one of ordinary skill in the art at the time the invention was made to include highlighting in the display of the original recording at least one user-selected phonetic unit, as in Hejna, Jr., so that the user can easily keep track of his or her selection visually.

Yamazaki teaches: (i) correcting elements of a text-to-speech segment dataset of parameters corresponding to a segment of the synthesized audio identified to be problematic (col. 25, line 61 – col. 26, line 16, teaches operations for changing [correcting] a velocity, pitch, phoneme, label, and/or voice tone setting);



(j) generating a new synthesized waveform corresponding to one or more adjusted parameters (see for example, Figs. 32 and 33, particularly elements 25E, E1, E11 and E12, which illustrate new synthesized waveforms corresponding to a velocity adjustment, as described in col. 25, ll. 18-39 and 61-67); and

(k) repeating steps (b) – (j) until a desired synthesized output is generated (see Fig. 24; col. 25, ll. 17-22, teaches “if it is determined...that an operation for terminating the processing for making new voice-generating information has not been executed, and at the same time that an operation for changing any parameter has not been executed, the processing...is repeatedly executed”; see also col. 25, ll. 53-60).

Regarding claim 26, Yamazaki teaches wherein the parameter updates and segment dataset corrections are applied in regenerating the synthesized audio (see for example, Figs. 32, 33 and 37, particularly elements 25E, E1, E11 and E12, which illustrate new synthesized waveforms corresponding to a velocity adjustment, as described in col. 25, ll. 18-39 and 61-67; see col. 25, ll. 40-52, “new filing”; col. 25, line 61 – col. 26, line 28, “change of velocity...change of pitch...for each label” col. 26, ll. 29-53, “change of parameters in...original synthesized waveform...object for editing”; and col. 23, ll. 1-16, “making new voice-generating information”; the information is a set of data for the voice segment (segment dataset), and is used to display the information in Fig. 33, and edits the information and the corresponding display).

Regarding claims 2 and 14, Yamazaki teaches [displaying] the parameters responsive to a user selection of at least a portion of the waveform, the displayed parameters correlating to the selected portion of the waveform (col. 26, ll. 29-47, “editing...basically the same processing...file as an object for editing is selected...treated as an original waveform,” if a user selects a file then he/she selects the entire waveform, at least a portion of the waveform, which is put on the display with the pitch, velocity, etc.).

Regarding claims 3 and 15, Yamazaki teaches identifying a portion of the waveform responsive to a user selection of at least one of the parameters, the identified portion of the waveform correlating to the selected parameters (col. 25, ll. 18-39, “velocity adjustment...velocity E1 in a time zone for the phoneme of ‘ka’...subdivided to velocity E11”; col. 25, ll. 61-67, “value of pitch...for each label”; Figs. 31-33; e.g. choosing the velocity for one of the phonemes to be changed identifies the corresponding phoneme/portion so that that portion of the waveform can be changed).

Regarding claims 8 and 20, Yamazaki teaches wherein said adjusting step comprises at least one action selected from the group consisting of deleting a pitch mark, inserting a pitch mark, and repositioning a pitch mark by deleting a phonetic unit label, [adding] a phonetic unit label, modifying a phonetic unit label, and repositioning the phonetic unit boundaries (see col. 24, ll. 6-43, “deletion of a pitch label...adding a new label...movement”; and col. 25, line 63 – col. 26, line 8).

Regarding claims 9 and 21, Yamazaki teaches wherein [said displaying parameters step] further comprises the step of [displaying] a waveform from the original recording along with the phonetic unit (col. 23, ll. 8-16, “natural voice is inputted...original waveform is displayed”; Fig. 33).

Regarding claims 10 and 22, Yamazaki fails to explicitly teach, but Miyatake teaches wherein edits to the waveform adjust parameters in the segment dataset (Figs. 3-4, col. 4, ll. 20-36, “displayed characters are edited by the inputting means...inputting means...separate...from each other...pauses are created between”; inserting the pauses edits the waveform and adjusts the starting and ending positions [parameters] of the phonemes).

Therefore, it would have been obvious for one of ordinary skill in the art at the time the invention was made to modify the teaching elements of Yamazaki with Miyatake in order to synthesize a natural speech that is close to the way a human being speaks, as taught by Miyatake in col. 1, ll. 20-30.

### ***Conclusion***

9. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37

Art Unit: 2626


CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

10. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Eunice Ng whose telephone number is 571-272-2854. The examiner can normally be reached on Monday through Friday, 8:30 a.m. - 5:00 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on 571-272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

EN  
9/5/07



DAVID HUDSPETH  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 2600